

ESCAPE

The New York Times Position

We are currently living through a massive platformisation of the Web. Independent, decentralised content is moving to private distribution channels and proprietary formats. The Web was not architected to produce this platformisation. It was not, however, designed to prevent it either. The belief was that it would be *sufficient* to enable a distributed environment for it to stay a distributed environment. That assumption has been demonstrated false.

Because of this mistaken assumption, Web technology has not traditionally been built with considerations towards policy, monetisation, ownership over data or content, or legal frameworks. In the Web and Internet communities, our ethos of old prevented us from building in the guardrails to protect against the emergence of excessive centralisation. We failed to bridge technology to society.

For the Web to remain a “[for good](#)” we must design Web technologies taking into account the context in which they will be used, beyond the purely technical. The Web’s full potential can only be for everyone; the architecture of the Web is meaningless if it does not account for the social and economic structure of the Web.

These problems manifest in several ways:

- Content and service providers compete with platforms for ad revenue but also share their audiences with them.
- Content is increasingly surfaced and rendered actionable as an immediate component of search or voice results, which threatens the revenue streams of content creators.
- Developing a brand and establishing a context of trust with users is a massive investment and a major asset for content producers but this production is then commoditised in environments carrying another brand and in such a manner that all sources are presented in a decontextualised manner.

In the same way that [STRINT](#) laid out a roadmap for an Internet somewhat more resistant to pervasive monitoring, we need to outline the path to a Web that is resistant to centralised economic dominance.

This is particularly important when considering a change such as Web Packaging that will have a major impact both on the Web’s architecture and on how publishers integrate with content

aggregators. Web Packaging fluidifies aggregation, but fluidity has a tendency to accrue value to the already dominant.

Data Controllership

One topic that is fundamental to the data economy but is not reflected in platform concepts is that of data controllership. This absence regularly leads to confusion.

The modern Web involves sites built from a complex system of interconnected services. While often represented in the simplistic view of a connection between a user and a resource, a typical Web interaction involves numerous third-party intermediaries such as browsers, ISPs, CDNs, Web hosts, database hosts, analytics services, and many more. This specialisation is an unavoidable consequence of the complexity of modern technology.

Reusing GDPR terminology (but very similar notions are reflected in other legal frameworks such as the CCPA, as well as in the DNT work), a data controller is a party that determines the means and purposes of the data processing. By contrast a data processor is a party that only processes the data at the instruction of a controller and cannot reuse it for its own purposes secondary to the transaction.

User expectations are grounded in a simple model in which they only interact with the one first party, and from a privacy standpoint this maps to an ideal situation in which the first party is the only data controller and there are no additional hidden controllers listening in. Under this view a CDN acting as data controller would be far more invasive than a third-party analytics script acting as a processor, entirely on behalf of the first party (which places strong constraints on data leakage and unexpected tracking).

While technical standards cannot make legal prescriptions they can convey the intended use of the technology and render explicit how the technological stack is meant to integrate within legal frameworks.

Since a primary use case for Web Packaging is to have a third party serve the content of the first party, the standard needs to indicate that the serving party, since it acts as a cache and not a replacement for the first party, must act as a data processor and only ever process data obtained from serving packages in a [permanently deidentified form](#).

Content Aggregation

Content aggregators are increasingly using content in new ways and in new products that go beyond the scope of search indexing. Moving forward we need to help publishers understand

how their content is being used and afford them greater control over their content and monetisation opportunities within these new environments.

Compatibility with Non-Advertising Business Models

Content on the Web can be monetised in a variety of manners. Some involve selling copies of a work (eg. books) while others require subscriptions (eg. newspapers). Some models are mixed in that some amount of content (a few articles per time period, the first chapter) can be provided free but beyond a certain level payment must be effected.

It is vitally important that Web Packaging does not prevent publishers from defining their own business models. Notably, it must remain possible to include only partial content in a package in all circumstances and it must be possible to integrate with online checks of eligibility.

It is worth noting that failure to do so would strongly incentivise demand for DRM support in packages, which has traditionally been seen as something to avoid.

Evidence-Based Standards

Web Packaging has other uses, but it is primarily seen by a large proportion of its stakeholders as a solution to problems that AMP created. Before we agree to solve those issues, should we not ask if AMP was a useful approach in the first place — and useful to whom?

Under what evidence criteria do we know that AMP has been successful? Its incentivisation through the carousel is widely recognised as a primary driver, but this does not entail that it would deliver benefits were that to stop.

If Google Search incentivised performance — which Google has often laid out as a next logical step — instead of a specific format would the problems we are seeking to address here still exist? If they did, would they be as pressing? Similar questions apply to other aggregation formats that are also being incentivised through unrelated means (a social network, an app store).

Given how substantial this architectural change is, it would be desirable to spend at least some effort in ensuring that we are addressing an actual need rather than merely giving an imprimatur to an approach whose value has not been independently appraised.

Robin Berjon
The New York Times Company